·全科医生方法与工具·

肺部听诊音数据库建库技术及方法研究

张冬莹[1，2]，叶培韬[3]，李洽胜[2]，简文华[2]，梁振宇[2]，郑劲平[2*]

1.999078 澳门特别行政区，澳门科技大学医学院
2.510120 广东省广州市，广州医科大学附属第一医院 广州呼吸健康研究院 国家呼吸系统疾病临床医学研究中心
3.510310 广东省广州市，广东省第二人民医院
[*]通信作者：郑劲平，主任医师/教授/博士生导师；E-mail：18928868238@163.com

【摘要】　当前无论是物理听诊器亦或是电子听诊器的肺音听诊结果仍然主要是依靠医生专业的听诊鉴别能力，尚未能够实现智能诊断判读。当患者在家受到肺部疾病影响时，无法自行发现肺部异常而耽误治疗；当处于呼吸道传染病救治过程中，入耳式的听诊器容易受到污染造成院内感染。尽管听诊音包含了丰富的健康状态信息，由于缺乏标准化的采集方法、分类标准和分析工具，使得听诊音的客观分析和应用在实践中受到了限制。本研究通过采用统一的听诊音采集设备和流程进行肺部听诊音数据采集、整理、数据库设计，使用软件 MatlabR2017a 进行数据管理和分析，建立了健康群体和肺部疾病患者群体的肺部听诊音数据库，制订一套标准的听诊音分类、标注规范、音频特征信号参数，构建一个用于存储、管理和分析肺部听诊音数据的系统，为肺部疾病的筛查、监测以及医学人工智能应用转化等相关研究提供重要的数据支持。研究积累了肺部听诊音音频数据库建库经验，为音频类数据库管理和分析提供有益的参考和借鉴，为支持后续医学人工智能辅助听诊应用于肺部疾病筛查与监测奠定基础，具有重要的医学价值和实际应用意义。

【关键词】　肺疾病；肺部听诊音；音频数据库；支持向量机；特征识别；数据分析
【中图分类号】　R 563　【文献标识码】　A　DOI：10.12114/j.issn.1007-9572.2023.0863

Study of Techniques and Methods for Building a Database of Lung Auscultation Sounds

ZHANG Dongying[1，2]，YE Peitao[3]，LI Qiasheng[2]，JIAN Wenhua[2]，LIANG Zhenyu[2]，ZHENG Jinping[2*]
1.Faculty of Medicine，Macau University of Science and Technology，Macau 999078，China
2.The First Affiliated Hospital of Guangzhou Medical University/Guangzhou Institute of Respiratory Health/National Clinical Research Center for Respiratory Diseases，Guangzhou 510120，China
3.Guangdong Second Provincial General Hospital，Guangzhou 510310，China
[*]*Corresponding author：ZHENG Jingping，Chief physician/Professor/Doctoral supervisor；E-mail：18928868238@163.com*

【Abstract】　Currently，the results of lung sound auscultation with either physical or electronic stethoscopes still rely mainly on the doctor's professional auscultation identification ability，which has not yet been able to realise intelligent diagnosis and interpretation. When patients are affected by lung diseases at home，they are unable to detect lung abnormalities on their own and delay treatment；when they are in the process of rescue and treatment of respiratory infectious diseases，in-ear stethoscopes are easily contaminated and cause nosocomial infections. Although stethoscopic sounds contain a wealth of information about health status，the lack of standardised collection methods，classification criteria and analysis tools has limited the objective analysis and application of stethoscopic sounds in practice.In this study，the data collection，arrangement and database design of the lung auscultation sound were carried out by using the unified auscultation sound collection equipment and process.The study used the software metlabR2017a for data management and analysis to create a database of lung auscultation sounds in a healthy group and a group of patients with lung disease.A database of lung auscultation sounds was established for healthy groups and groups of patients with lung diseases.A standard set of classification of auscultatory tones，labelling specifications，audio characteristic

基金项目：澳门科技大学发展基金项目（0070/2020/A2）
引用本文：张冬莹，叶培韬，李洽胜，等.肺部听诊音数据库建库技术及方法研究［J］.中国全科医学，2024. DOI：10.12114/j.issn.1007-9572.2023.0863.［Epub ahead of print］［www.chinagp.net］
ZHANG D Y，YE P T，LI Q S，et al. Study of techniques and methods for building a database of lung auscultation sounds［J］. Chinese General Practice，2024.［Epub ahead of print］.

signal parameters were developed.Building a system for storing，managing and analysing lung auscultation sound data to provide important data support for research related to the screening and monitoring of lung diseases and the translation of medical artificial intelligence applications.The study accumulated the experience of building an audio database of lung auscultation sounds，provided a useful reference for the management and analysis of the audio database，and laied the foundation for supporting the subsequent application of medical artificial intelligence-assisted auscultation in the screening and monitoring of lung diseases，which was of great medical value and practical application.

【Key words】 Lung diseases；Lung auscultation sound；Audio database；Support vector machine；Feature recognition；Data analysis

随着医学科学技术的发展，肺部疾病的诊断越来越依赖于医学数据，肺部听诊音作为肺部疾病诊断的重要依据之一，其标准音频数据库建立和数据分析显得尤为重要。然而，建立一个高质量的肺部听诊音数据库并非易事。医务工作者需要考虑如何确保数据的准确性和可靠性，如何有效地组织和存储数据，如何利用人工智能和机器学习技术对数据进行处理和分析等问题[1]。近年来，人工智能和机器学习技术为解决这些问题提供了新的可能。这些技术能够提高数据采集、处理和分析的效率和准确性，使人们从海量的医疗数据中提取出更有价值的信息[2]。因此，利用这些技术构建肺部听诊音数据库系统，成为当前肺部听诊音数据库建库技术及方法研究的重要方向。

肺部听诊是医生在诊断和监测呼吸系统疾病时常用的重要手段。肺部音频信号可以提供许多有价值的信息，如呼吸音、杂音等，可以辅助医生诊断和判断疾病的类型和严重程度[3]。然而，肺部听诊音的解读和分析常需要医生具备丰富的经验和专业知识。传统的肺部听诊音主要依靠临床实践和经验积累，但这种方式的局限性在于其主观性较强，存在个体差异，对医生培训和病情判断的一致性有一定影响[4，5]。为了解决这些问题，国内外研究者探索利用机器学习和信号处理技术来对肺部听诊音数据进行分析和建模。通过机器学习和神经网络模型，AI 可以对大量经过声音标注的呼吸音数据进行快速学习，从而实现对呼吸音的自动分析[6]。呼吸音的自动识别和分类可以分为两步：特征提取和识别分类。然而，在呼吸音领域的 AI 发展过程中，面临一个主要问题，即缺乏可用于研究算法和比较不同算法的公共高质量大型呼吸音数据库，这也是 AI 在呼吸音领域发展的主要障碍[7]。建立肺部听诊音数据库可以为自动化诊断系统、疾病监测和健康评估等提供基础。随着智能医疗和远程医疗的发展，建立一个公共的、大规模的肺部听诊音数据库也具有重要意义。这样的数据库可以促进医疗资源共享和协作研究，提高肺部疾病的诊断和治疗水平。因此，对于肺部听诊音数据建库技术及方法的研究可以帮助改善传统肺部听诊方法的不足且推动肺部疾病的早期诊断和治疗。

1 肺部听诊音的采集

1.1 音频数据采集设备

本研究采用美国 3M 公司的 3M^{TM} Littmann®3200 听诊器型电子听诊器获取肺部听诊音信息数据，该听诊器有以下特点：通过数字信号传导而非声学振动来采集呼吸音数据；可以收集更广的音频范围，并进行数字信号转换；可同步生成视觉声音图谱，即在听到呼吸音的同时看到呼吸音的波形图谱。上述特点有助于使用者调整听诊技术，获得更准确的听诊结果[8]。该类型听诊器与普通听诊探头相比，可以听到更低频的声音，音域更广、响度更大，且具有良好的环境降噪效果。譬如其采集频率范围为 20~2 000 Hz，采样频率为 4 kHz，可将声音放大 24 倍，并能降低听诊探头端的杂音，提高声学传导质量。此外，其还具备录音、存储和数据传输功能，可以连接电脑，并可使用匹配的音频软件进行分析，是建立标准肺部听诊音的理想工具。

1.2 数据获取

1.2.1 研究对象准备：本研究研究对象均来自广州医科大学附属第一医院（患者来源于呼吸科住院部，健康群体来源于医院员工），入组前使其充分知情肺部听诊音建库研究，告知采集流程和需其配合的环节，并请其签署知情同意书面材料。依据被采集人的年龄、性别、基础病、被采集时的身体状态等记录详细信息并做好分组。本项目是一项开放式的研究探索，参考前人相关研究，招募总计 350 名参与者作为研究样本，这些参与者将被划分为各含 50 人的七组，以便于进行详尽的比较和分析。该设计标准可确保研究的准确性和有效性，同时也保护参与者的权益。本研究经广州医科大学附属第一医院伦理审查通过（医科伦审 2017 第 82 号）。

1.2.1.1 纳入标准：（1）志愿参与。（2）年龄≥ 18 岁。（3）健康人群：①既往无肺部慢性疾病、无粉尘常期暴露史且无吸烟习惯，不受器质性心血管疾病的困扰者；②近 3 周内无急性上呼吸道感染及咳嗽、咳痰等呼吸道症状；③所有研究对象在过去六个月内进行的健康体检中，胸部 X 光未发现任何肺部病变，并且心电图表现正常的。COPD 患者：①有呼吸困难、慢性咳嗽或咳

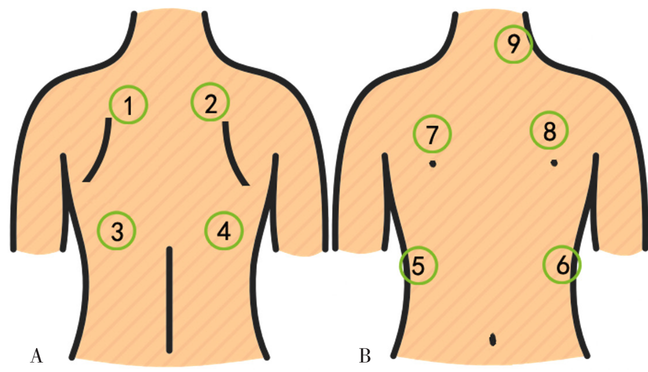
痰，有危险因素暴露史；②肺功能检查，吸入支气管扩张剂后第一秒用力呼气量（FEV1）/用力肺活量（FVC）<0.7；③无大量胸腔积液或气胸的并发症；④临床上确诊为COPD；符合第1~3项或符合第3~4项。支气管哮喘患者：①经常出现反复喘息、气急、胸闷或咳嗽，多与接触变应原、冷空气、物理、化学刺激、病毒性上呼吸道感染、运动等有关；②发作时在双肺可闻及散在或弥漫性、以呼气相为主的哮鸣音，呼气相延长；③上述症状通常可经使用平喘药物治疗后缓解或自行缓解；④除外其他疾病所引起的喘息、气急、胸闷或咳嗽；⑤对症状不典型者（如无明显喘息或体征），应最少具备以下三项中至少一项阳性：（a）支气管激发试验（或运动激发试验）阳性，（b）支气管舒张试验阳性，（c）昼夜PEF变异率≥20%；⑥临床上确诊为支气管哮喘；符合第1~4项或符合第4~5项或符合第6项者。支气管扩张患者：①幼年时期有诱发支气管扩张的呼吸道感染史，如麻疹、百日咳或流感后肺炎病史，或肺结核病史等；②出现长期慢性咳嗽、咳脓痰或反复咯血症状；③高分辨率计算机断层扫描（HRCT）显示支气管扩张的异常影像学改变；④临床上确诊为支气管扩张；符合第1~3项或符合第4项者。间质性肺疾病患者：①在双侧下肺区会有干啰音或湿啰音，尤其是在吸气末期更加明显；②胸部X光显示出弥漫性的阴影或不常见的结节样间质性阴影，以及支气管壁增厚和血管周围间隙的突出，小的规则或不规则阴影以及小的外周性环状阴影；③肺功能检测揭示了限制性通气功能障碍和气体交换功能减退；④通过对支气管肺泡灌洗液（BALF）的细胞学、病原学、生化和炎症介质等检测，可以依据BALF中炎症免疫效应细胞的比例将间质性肺疾病分为淋巴细胞增多型和中性粒细胞增多型；⑤通过支气管肺活检或外科肺活检获取肺组织进行病理学检查确诊为间质性肺疾病；⑥临床上确诊为间质性肺疾病。肺炎患者：①咳嗽、咳痰，或原有呼吸道疾病症状加重，并出现脓性痰，伴或不伴胸痛；②发热，体温≥38 ℃；③在肺部实变病变或湿啰音中可以观察到感染的迹象；④白细胞计数（WBC）>10×10^9/L或4<×10^9/L，伴随或不伴随核左移；⑤胸部X射线检查可能会显示斑片状浸润阴影或间质性变化，时伴随有胸腔积液；⑥临床上确诊为肺炎；符合第1~4项中任意一项和第5项。肺部肿瘤患者：①胸部CT提示肺部肿瘤，或提示肺部占位且考虑肿瘤可能性大；②肺部肿瘤直径>2 cm；③临床上考虑肺部肿瘤可能；④未合并大量胸腔积液和气胸；⑤未行手术切除；⑥临床上已确诊为肺部肿瘤；符合第1~5项或符合第项2、4、5、6项者。肺部肿瘤患者：①胸部CT提示肺部肿瘤，或提示肺部占位且考虑肿瘤可能性大；②肺部肿瘤直径>2 cm；③临床上考虑肺部肿瘤可能；④未合并大量胸腔积液和气胸；⑤未行手术切除；⑥临床上已确诊为肺部肿瘤；符合第1~5项或符合第项2、4、5、6项者。

1.2.1.2 排除标准：（1）确诊患有癫痫并且目前正在接受药物治疗以控制病情；（2）妊娠；（3）肺部切除手术的患者；（4）处于昏迷状态或有意识障碍的患者；（5）无法配合研究流程的个体；（6）在重症监护病房（ICU）接受治疗或病情处于危急状态的患者；（7）未签署知情同意书；（8）同时满足以上两个或两个以上的纳入标准的患者。

1.2.2 录音环境准备：要求独立安静、温度适宜的空间进行肺部听诊录音，录音时环境噪音不高于40~50 dB。

1.2.3 资料收集：收集研究对象的病史详细信息及其相关检查资料（包括年龄、身高、体质量、主要症状、吸烟史、呼吸系统相关主要诊断、肺CT结果、肺功能检查报告数据等）。

1.2.4 录音方法：嘱研究对象放松心情，平静呼吸数分钟。选用3M™ Littmann®3200型电子听诊器，按人民卫生出版社第9版《诊断学》肺部听诊要求对喉部、左右上、中、下肺部、左右肩胛骨下共9个部位进行听诊（图1），录取所有入组研究对象的肺部听诊音，同时填写资料表格。每个点至少采集15秒的呼吸音，研究对象保持安静。听诊器探头紧贴研究对象皮肤，开始采集后，嘱研究对象深呼吸，确保每次采集的呼吸音包含2~3个完整的呼吸周期[9]。



注：A表示背面，B表示正面。

图1 肺部听诊音采集部位示意图

Figure 1 Schematic diagram of the site of collection of lung auscultation sounds

1.2.5 质量控制：由2名呼吸内科主治医师分别听取所采取的肺音进行录音鉴别与诊断，如存在分歧，再组织第3名呼吸内科主治医师听取鉴别，消除歧义后录用数据。

2 数据传输与存储

2.1 数据传输

3M™ Littmann®3200 型电子听诊器内存有限，可存储 12 段长度为 30s 的音频，因此录制完每个研究对象都需将肺音音频通过蓝牙和听诊器适配的软件 StatAssist 传输到电脑端，需注明每段音频来源于具体研究对象所对应的肺部部位。

2.2 数据存储

传输到电脑端的肺音音频以 zsa 格式保存，文件名会自动产生一串编码，为了与相应研究对象对应，可在编码后加上研究对象姓名以对应到具体的人，该段 zsa 格式的音频即为原始数据。

2.3 音频导出

将 zsa 格式音频导出为 wav 格式保存，以便于后续音频标注和分析。将同一研究对象的肺音音频导入同一文件夹，文件夹以原文件名命名，按每段音频所对应部位重命名，以便于每条肺音可以准确对应到某个研究对象的具体身体部位。

2.4 数据脱敏

包含研究对象姓名的肺音音频数据属于敏感信息，可能因设备遗失或电脑病毒而导致信息泄露，故需对数据进行脱敏。将文件名全部随机重命名，并将重命名后的数据与研究对象姓名、基本信息等对应起来。

2.5 云端数据库的建立

存储在电脑端的数据有可能因电脑故障而丢失，故将脱敏数据传入云端是很有必要。设立一个只供本地网络访问的云端数据库，对数据的存储及安全都有良好的保障。将未标注、标注后及审核后各阶段数据分别保存在云端数据库，方便统计数据纳排情况及数据溯源。

3 音频数据的预处理

3.1 肺音降噪

肺音信号的预处理首先是去除肺音中的噪声，目的是提高呼吸音信号的清晰度和质量，准确地捕捉和识别肺部声音信号。电子听诊器虽然有一定的降噪能力，但在肺音的采集过程中会不可避免地引入一些噪声，包括来源于外部环境杂声、听诊器电路内部产生的电流干扰、以及探头与皮肤接触面的摩擦声和心音等人体内部噪声等[10]。原始肺音信号的噪声来源很多，如果不能有效地滤除噪声，将对肺音的识别过程产生较大影响。

3.1.1 肺音信号的高通滤波去噪：肺音信号的高通滤波去噪的目的是去除低频噪音，保留高频部分，从而提取出肺音信号的有用信息。

肺音信号中的呼吸音和心脏音一般集中在较高的频率范围内，而噪音通常是低频信号。通过应用高通滤波器，可以抑制低频噪音的干扰，使得肺音信号更加清晰。高通滤波的关键是选择合适的截止频率和滤波器参数，以充分去除低频噪音，同时保留肺音信号中的有用信息。

3.1.2 肺音信号的低通滤波去噪：根据肺音信号的信号特点，设计一个低通滤波器，可以将肺音信号中高频部分去除，保留低频部分。进行滤波操作，将设计好的低通滤波器应用到肺音信号上，通过滤波操作可以将高频噪声去除，同时保留肺音信号中的低频成分。进行信号重构，经过滤波操作后，得到的信号可能会有一定的变形，需要进行信号重构处理，使其恢复到原始形态。

3.2 人工标注

3.2.1 音频标注软件：Audacity 是一款公开免费的音频处理软件，用于录音和编辑音频，是自由、开放源代码的软件分析。该软件具有的可视化工具可用于分析音频或其他信号数据，支持 Python，Perl 或任何支持命名管道的语言编写脚本，兼容性强，可在 Mac OS X、Microsoft Windows、GNU/Linux 和其他操作系统上运作。

3.2.2 肺部听诊音标注者的选择：肺部听诊音标注的准确性在很大程度上影响模型的识别效率，呼吸内科医生对肺部听诊音的识别比较有经验，也有利于提高音频数据标注的准确度，故标注者的选择建议是呼吸内科医生。由于采录设备录制的声音与临床实际听到的声音有些许差别，所以每位标注者需经过专业培训后方能正式参与肺音标注。

3.2.3 标注流程：将待标注音频导入 Audacity 软件，将默认的波形图转换为频谱图模式，频谱图可直观显示一个呼吸周期及该呼吸周期的吸气相和呼气相，有利于快速选区并标注，选定区域后输入标签。为了使标注更加精确，将整段音频中每个呼吸周期先框选出来，再对吸气相和呼气相进行区分后作定性（标注为正常、湿啰音、哮鸣音、喘鸣音或鼾音），有助于呼吸音的特征提取。不完整的呼吸片段可能会降低特征提取的效果，故只标注较为完整的呼吸周期。

标注的内容包括：呼吸周期的呼气段和吸气段，正常、湿啰音和干啰音发生的位置（精确到 ms），干啰音细分的类型。还包括对呼吸音音频质量等级标注，音频质量等级按照音频质量的好坏，依次分级为Ⅰ、Ⅱ、Ⅲ三个等级，其中Ⅰ级最优。评价标注包括音频的信噪比、噪音时长占比及人工主观评价。标注完成后，带标签的音频将以 Audacity Project File（aup3）格式保存。标注时所有标注者使用同种品牌同种型号的收听设备（铁三角 ATH-M20X 头戴式耳机），使所有标注者尽可能获得较为一致的听觉体验。

3.2.4 标签审核：对于标注过程中有争议的部分，需进行公开讨论，最终获得一致意见的，予以采纳；若最终无法达成一致意见，则剔除该条数据。全部音频标注完成后，由呼吸与危重症学科高年资医生对所有音频进行审核，修改错误或不规范标签，并剔除低质量数据（噪音太多影响判断或音量太小难以分辨的音频）。经过审

核的标签即可用于训练 AI 模型。

3.3 肺音识别机器学习算法及其模型建立过程

3.3.1 湿啰音和哮鸣音识别及其持续时间：湿啰音和哮鸣音是 2 种常见的呼吸音，可以通过听诊器或其他听诊设备进行识别。湿啰音是由于呼吸道黏膜分泌物增多导致的湿润声音。其通常类似于呼吸道中的液体在被吸入或呼出时产生的咕噜声。湿啰音可以是细小而清澈的，也可以是较为粗糙和响亮的。它们可能由于感染、充血、分泌物过多或阻塞引起。湿啰音的持续时间可以因患者情况而异，从持续数秒到几分钟不等[11]。哮鸣音是由于气流通过狭窄的呼吸道导致的高音频咳嗽声或呼哧声。这种声音通常类似于从狭窄的空间经过的高速气流所产生的噪音。哮鸣音可以是连续的或间断的，并且强度和音调可能会变化；其可能由于肺部狭窄、支气管痉挛、梗阻或病变引起。哮鸣音的持续时间也可以因患者情况而异，从数秒钟到数分钟不等[12]。

湿啰音是偶发性的非连续信号，基于音频数据分析，笔者发现其持续时间一般 <20 ms。本研究团队使用软件 Matlab，呈现出 16 个湿啰音频谱图，其中红色的部分为以湿啰音峰值为中心的 20 ms 信号，详见图 2。

哮鸣音一般布满整个呼气相或吸气相，经过大量数据的统计，其持续时间一般 >250 ms。本研究采用红色标出以哮鸣音片段峰值为中心的 500 ms 信号，详见图 3。经过大量数据的测试，本研究团队发现 500 ms 在绝大多数情况下能够覆盖呼吸相或吸气相的 80% 以上。

3.3.2 湿啰音的波峰和哮鸣音的波峰定量分析：通过对超过 1500 个湿啰音进行统计，发现湿啰音的波峰幅值是湿啰音平均幅值的 2 倍以上。使用软件 Matlab，呈现出了 16 个湿啰音频谱图，其中红色的线是湿啰音平均幅值的 2 倍，详见图 4。

通过对超过 100 个哮鸣音测试，发现用 200 ms 长的汉明窗对哮鸣音能量值进行低通滤波，绝大部分低通滤波后的哮鸣音片段有 160 ms 的峰值部分会大于滤波后信号的平均值（图 5 的红线）。如图 5 所示，本研究用该方法找到了该哮鸣音的所有哮鸣音片段峰值，并用绿色线标出。

3.3.3 利用支持向量机（support vector machine，SVM）法进行湿啰音和哮鸣音的识别：SVM 是一种常用的监督学习算法，主要用于分类和回归分析。它是一种二类分类模型，通过最大化分类边界与数据之间的间隔来对

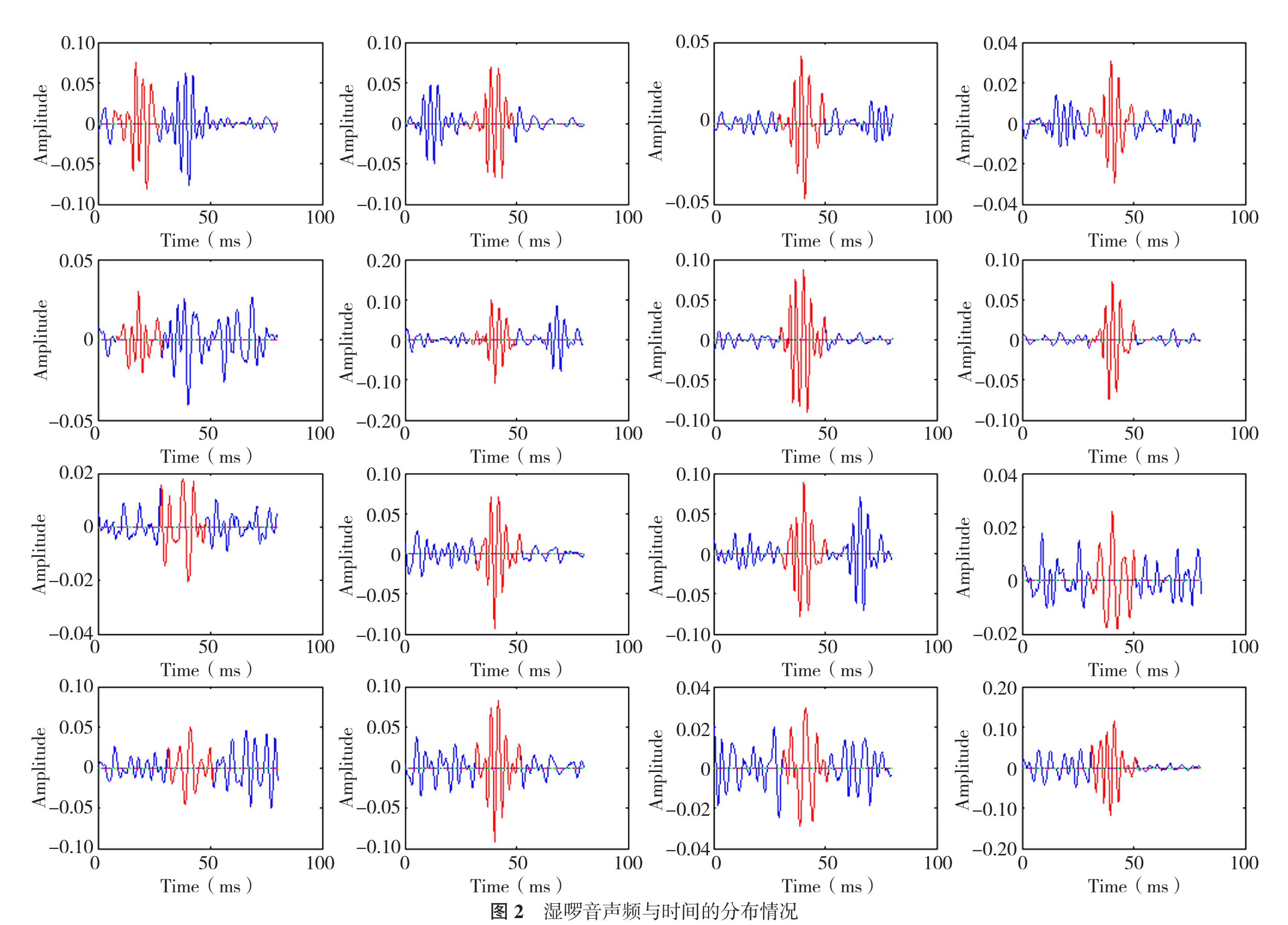


图 2 湿啰音声频与时间的分布情况

Figure 2 Distribution of the frequency and time of moist rale

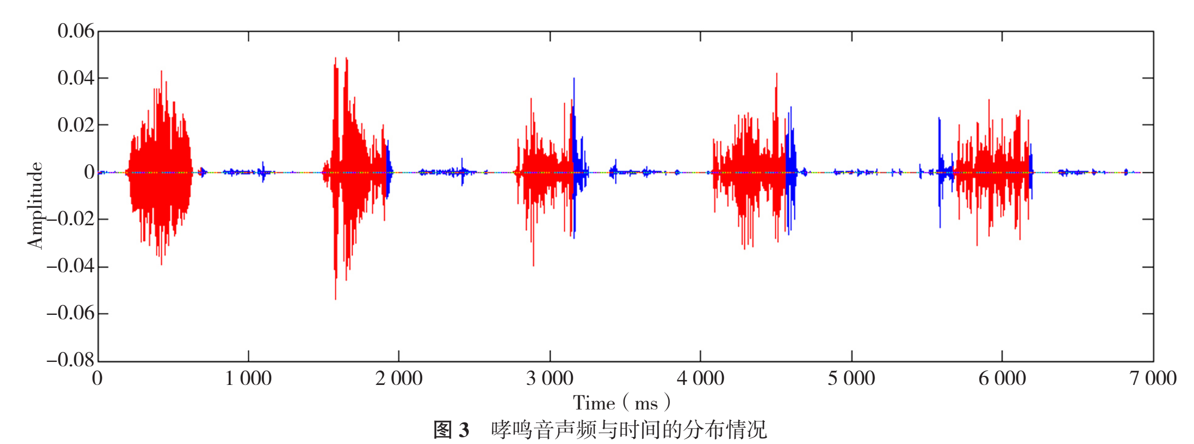
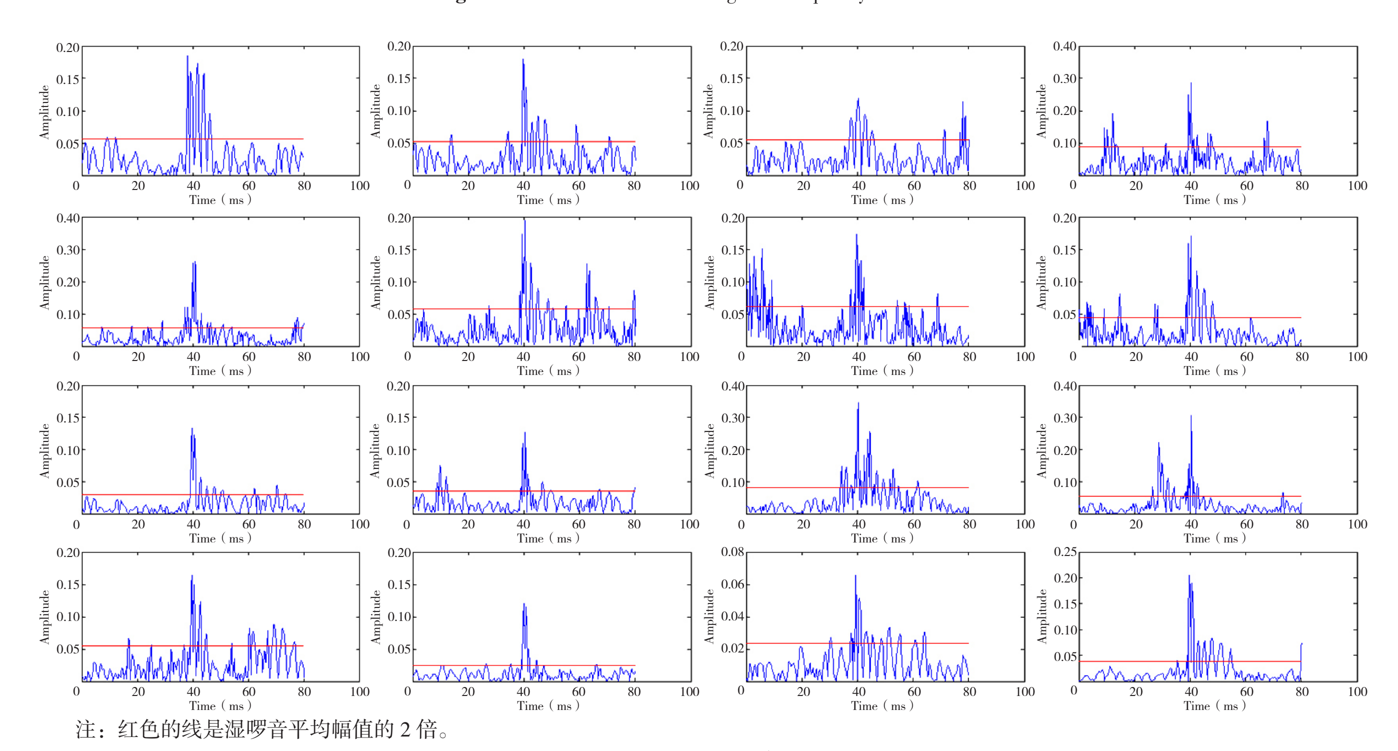


图 3 哮鸣音声频与时间的分布情况

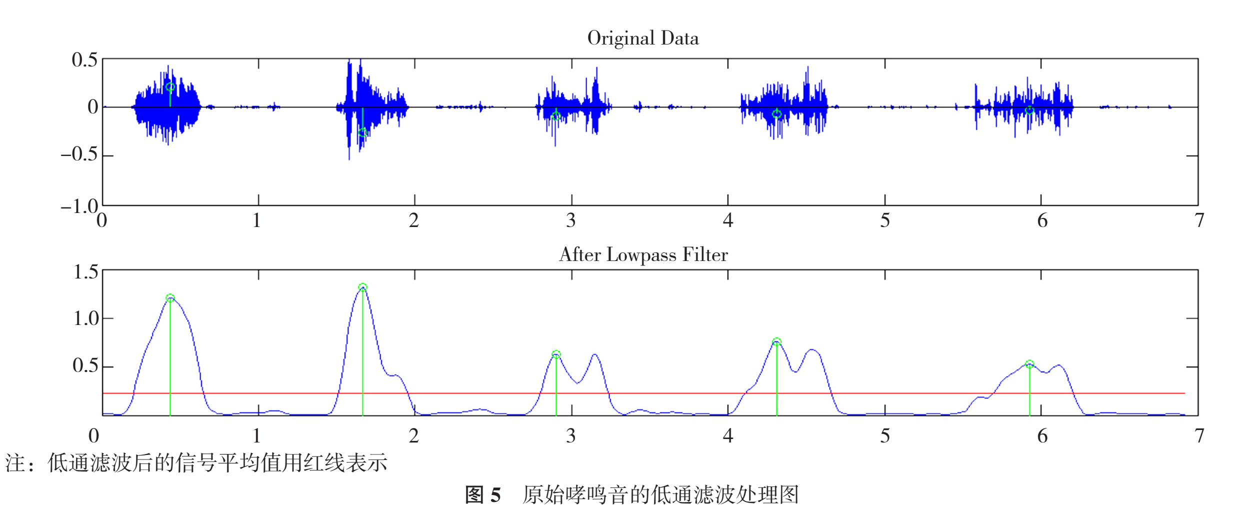
Figure 3 Distribution of wheezing soud frequency versus time



注：红色的线是湿啰音平均幅值的 2 倍。

图 4 湿啰音频谱图

Figure 4 Spectrogram of moist rale



注：低通滤波后的信号平均值用红线表示

图 5 原始哮鸣音的低通滤波处理图

Figure 5 Low-pass filtering of the original wheezing soud

新的数据进行预测。SVM 是基于统计学习理论，通过在高维空间中找到最优分类超平面来对样本进行分类，还可以用于回归分析、异常检测和时间序列预测等任务，是一种广泛应用于机器学习和数据挖掘领域的算法，具有较高的准确性和鲁棒性[13]。

利用 SVM 进行肺部听诊音识别是一种基于机器学习和信号处理的技术，其基本原理是通过对肺部听诊音信号进行特征提取和分类，从而实现肺部听诊音的识别和分类。通过训练 SVM 模型，可以根据已有的肺部听诊音数据和对应的标签（如正常肺音、异常肺音）来建立分类模型[14]。这个模型可以用于对肺部听诊音数据进行预测和分类，从而辅助临床诊断。SVM 在肺部听诊音识别中的优势在于，其可以处理高维数据和复杂的非线性关系，通过核函数将原始数据映射到高维特征空间中，从而避免了手动设计特征和选择特征空间的复杂性。同时，SVM 具有较高的准确性和鲁棒性，能够处理各种不同类型和质量的肺部听诊音数据。

在本研究中，笔者团队尝试利用 SVM 法对肺部听诊音进行湿啰音和哮鸣音的识别，主要包括音频数据库的建立、信号特征提取、向量机训练及向量机预测 4 个部分。与绝大多数研究不同的是，本研究的训练数据集包含的并不是整个肺音采录文件，而是经过预处理的音频数据片段，包括正常音频片段和异常音频片段。其中，湿啰音片段以湿啰音片段峰值为中心的20 ms信号组成，哮鸣音片段以哮鸣音片段峰值为中心的 500 ms 信号组成。这样做的好处主要有两个，其一，以峰值为中心的信号片段信号强，信噪比高，从而可以比较好的降低背景噪音造成的干扰，提高算法的准确度；其二，以峰值为中心的信号片段数据量较小，从而明显减少信号处理时间，提高算法的实时性、实用性。

在采集肺部听诊音的过程中，本研究发现听诊器探头薄膜随着健康受试者吸气胸腔扩张被动错位或探头在转移过程中都会产生伪湿啰音，且现阶段无法避免[15]。为了能够提高识别湿啰音的准确度，笔者也对伪湿啰音进行人工标注和识别且建立伪湿啰音的音频数据库，为后续能够实现对伪湿啰音的自动识别作准备。本研究提取的每个湿啰音片段时长为 20 ms，通过算法及半人工参与的方式建立真性和伪湿啰音片段库，详见图 6，显示的是部真、伪湿啰音片段建库过程。

为了进一步处理和分析肺音信号，本研究运用小波变换方法，对湿啰音与伪湿啰音进行处理，并提取其频域特征。这些特征值是从信号的基本波形和频谱中提取出来的。将计算出的时域和频域特征值作为输入变量，利用经过训练的 SVM 模型来对它们进行分类。通过信号的小波变换分解、特征值的计算，以及 SVM 分类模型的应用能够以较高的准确率来识别和区分正常与异常的呼吸音。具体的算法识别过程与结果如下：

（1）本研究结合自动算法和人工审核技术，深入分析肺音数据。研究建立了两个数据库，一个收集真实湿啰音，另一个收集伪湿啰音。由于真实湿啰音通常不超过 20 ms，研究确保所有声音样本长度为 20 ms，以精确捕获这些短暂信号。通过分析这些样本，提取了关键的时域和频域特征，这对分类湿啰音很重要，详见图 7。研究还使用 SVM 算法对特征进行训练和优化，成功区分真假湿啰音。

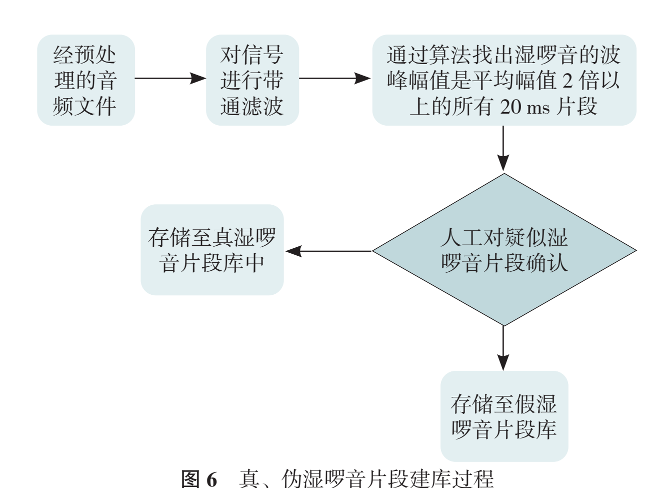


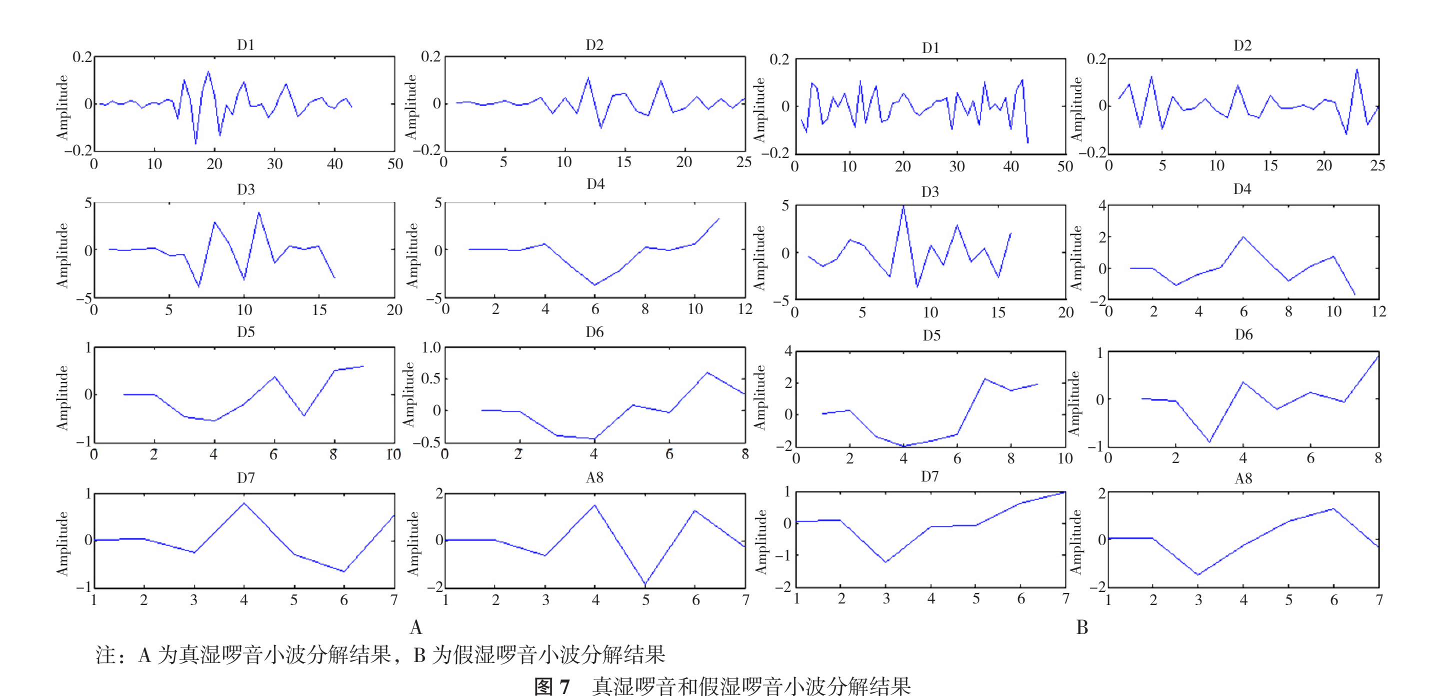
图 6 真、伪湿啰音片段建库过程

Figure 6 The process of creating audio databases of true and fake moist rale

（2）研究首先对肺音信号进行预处理，然后通过带通滤波器处理，强调 100~2 000 Hz 频率范围内的关键信号，该范围被认为含有诊断相关的重要信息。使用 20 ms 的时间窗口逐步扫描整个信号，以连续搜索并捕捉瞬时湿啰音。每发现一个潜在的湿啰音，就从该段信号提取时域和频域特征，并将这些特征用于 SVM 模型，以区分真实和伪湿啰音。

通过算法及半人工参与，从肺音中提取两个声音库，其中一个声音库由湿啰音组成，另外一个由伪湿啰音组成。由于湿啰音基本小于 20 ms，因此本研究提取的每个声音档为 20 ms。

提取真湿啰音和伪湿啰音的时域特征值及频域特诊值，然后利用向量机进行训练，得出向量机参数。利用该向量机参数可以很好的分辨真湿啰音和伪湿啰音。先进行信号预处理，重采样和识别弱信号，然后对呼吸音信号进行带通滤波（肺音信号主要在 100~2 000 Hz），然后计算智能阈值（每段呼吸音信号都有自己的阈值，根据该阈值找出潜在湿啰音），然后用 20 ms 的窗沿着肺音信号移动，找出所有潜在湿啰音。提取潜在湿啰音的时域特征和频域特征，并用向量机进一步判断该潜在湿啰音是真湿啰音还是伪湿啰音，最后根据该呼吸音平均 1 s 所含湿啰音的数量来判断该呼吸音是否为湿啰音



注：A 为真湿啰音小波分解结果，B 为假湿啰音小波分解结果

图 7 真湿啰音和假湿啰音小波分解结果

Figure 7 Results of wavelet decomposition of true moist rale and false moist rale

或哮鸣音，图中的蓝色圈圈标示的是伪湿啰音，红色圈圈标示的是真湿啰音，详见图 8。

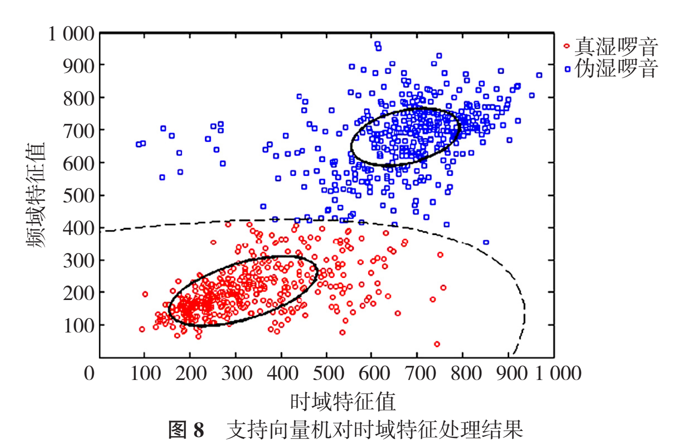


图 8 支持向量机对时域特征处理结果

Figure 8 Support vector machine results for time domain feature processin

由于哮鸣音会布满整个呼气相或吸气相，所以本研究提取的每个哮鸣音片段为 500 ms，绝大部分情况下，500 ms 能够覆盖整个呼气相或吸气相 80% 以上。通过算法及半人工参与的方式建立哮鸣音片段库，如图 9 所示。应用训练向量机识别湿啰音和哮鸣音的过程，详见图 10~11。应用向量机识别肺音的完整流程，详见图 12。

3.3.4 利用小波分解法对肺部听诊音进行特征提取：小波分解是一种信号分析方法，用于将信号分解为不同尺度和频率的成分，以便更好地分析信号的特点和特征。小波分解通过使用小波函数进行伸缩和平移操作，将信号分解为多个频率成分，譬如湿啰音分解后的信号频率分布，详见表 1。每个成分具有不同的尺度和形状。这种分解方法可以在一定程度上消除信号中的噪声，并且可以对信号进行压缩，减小数据量。小波分解在信号处理、图像处理、医学影像等领域有着广泛的应用[16]。

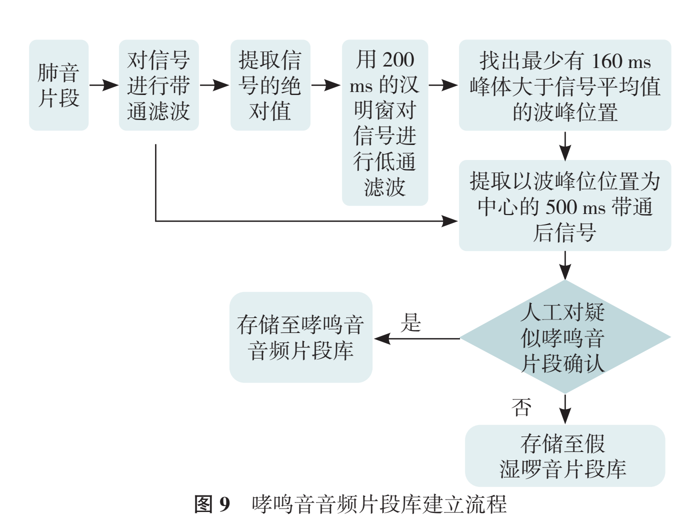


图 9 哮鸣音音频片段库建立流程

Figure 9 The process of creating an audio database of wheezing soud

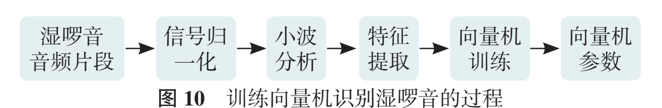


图 10 训练向量机识别湿啰音的过程

Figure 10 Training Vector Machines to recognise moist rale

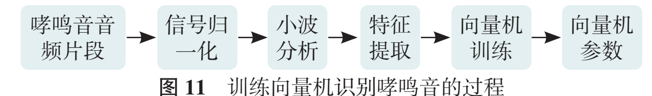


图 11 训练向量机识别哮鸣音的过程

Figure 11 Training Vector Machines to recognise wheezing soud

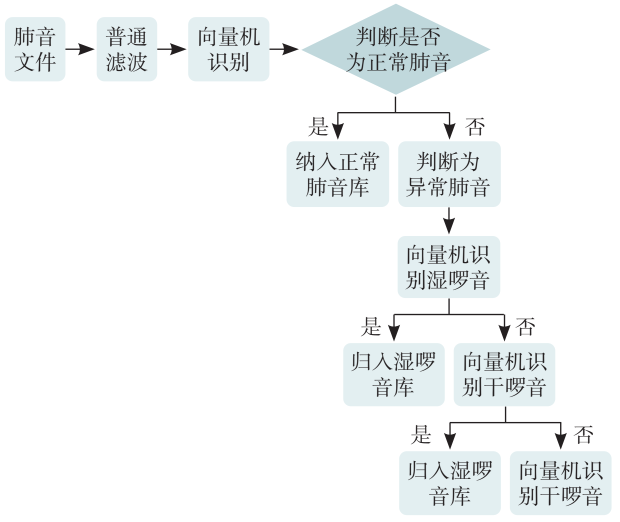
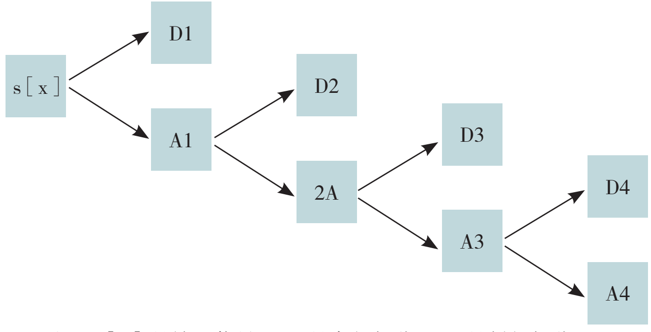


图 12　向量机识别肺音流程

Figure 12　Vector Machine Recognition of Lung Sounds Process

肺部听诊音是一种复杂的生理信号，包含了多种频率和特征的成分。小波分解应用于肺部听诊音特征提取的原理主要是基于信号的频域特性。小波分解通过使用小波基函数对肺部听诊音进行分解，将信号分解为不同尺度和频率的成分，从而更好地分析信号的特点和特征。这些成分可以反映肺部听诊音在不同时间和空间上的变化特征，从而为特征提取和分析提供更丰富的信息[17]。通过小波分解，可以提取肺部听诊音中的不同频率成分，例如高频成分可能代表呼吸急促、咳嗽等症状，低频成分可能代表呼吸节律等特征。这些特征可以用于构建分类模型，用于识别不同的肺部疾病或症状。

本研究将输入声音片段分解成和 D1、D2、D3、D4、D5、D6、D7、A8 共 8 段不同频率的信号，小波分解肺音的过程，详见图 13。湿啰音音频的分解结果，详见图 14。



注：s［x］是输入信号，Dn 是高频部分，An 是低频部分。

图 13　小波分解肺音的过程

Figure 13　Wavelet decomposition of lung sounds

对于分解出的 8 段信号，本研究提取如下特征（共提取 32 个特征值）：（1）提取每段信号绝对值的平均值，共 8 个平均值；（2）提取每段信号绝对值的方差，共 8 个方差；（3）提取每段信号绝对值的最大值，共 8 个最大值；（4）提取 8 个平均值相邻值得比值。分别为 mean(abs(s[x]))/mean(abs(D1), mean(abs(D1)) /mean (abs (D2)), mean (abs (D2))/mean (abs (D3)), mean (abs (D3)) /mean (abs (D4)), mean (abs (D4)) /mean (abs (D5)), mean (abs (D5))/mean (abs (D6)), mean (abs (D6)) /mean (abs (D7)), mean (abs (D7)) /mean （ abs （ A7 ））。

特征 1~8（8 个平均值）和特征 17~24（8 个最大值）代表了湿啰音在不同频率的能量分布，特征 9~16（8 个方差）和 25~32（相邻值得比值）代表了湿啰音在不同频率的能量变化情况。

表 1　湿啰音分解后的信号频率分布

Table 1　Frequency distribution of the signal after decomposition of moist rale

分解后的信号	频率分布（Hz）
D1	1 000~2 000
D2	500~1 000
D3	250~500
D4	125~250
D5	62.5~125.0
D6	31.2~62.5
D7	15.6~31.2
A8	0~15.6

4　应用与结论

应用上述采集肺部听诊音的操作流程与建立肺部听诊音音频数据库的技术方法，本研究采集了 392 例呼吸慢病患者，合格入库 3362 条音频数据，另有 56 例健康人的肺部听诊音 500 条音频数据，以及采用国立台湾大学生物医学电子与生物信息学研究生院公布的 HF_Lung_V1、国际生物医学与健康信息学会议公布的呼吸音数据库及上海交通大学公布的 SPRSound 数据库共 9469 条音频数据。应用软件 MetLab R2017a 进行数据管理和分析，对音频数据进行预处理、数据标注、向量机识别、特征提取等，基于呼吸音分析算法软件，可以自动导入呼吸音并频谱化显示，并通过 SVM 算法建立的正常和异常呼吸音模型库，自动识别出正常肺音，干、湿啰音，并在呼吸音频谱图上精确标注所在位置。应用程序界面识别肺部听诊音示例，详见图 15。

呼吸疾病的早发现、早诊断、早干预，有利于获得呼吸慢病患者有效管理的关键节点。肺部听诊音数据库的建立和相关方法研究为日后实现对肺部听诊音的智能判读和居家应用健康管理产品的研发奠定基，甚至可以协助呼吸道传染临床诊治，也为后期研制国产质优价廉

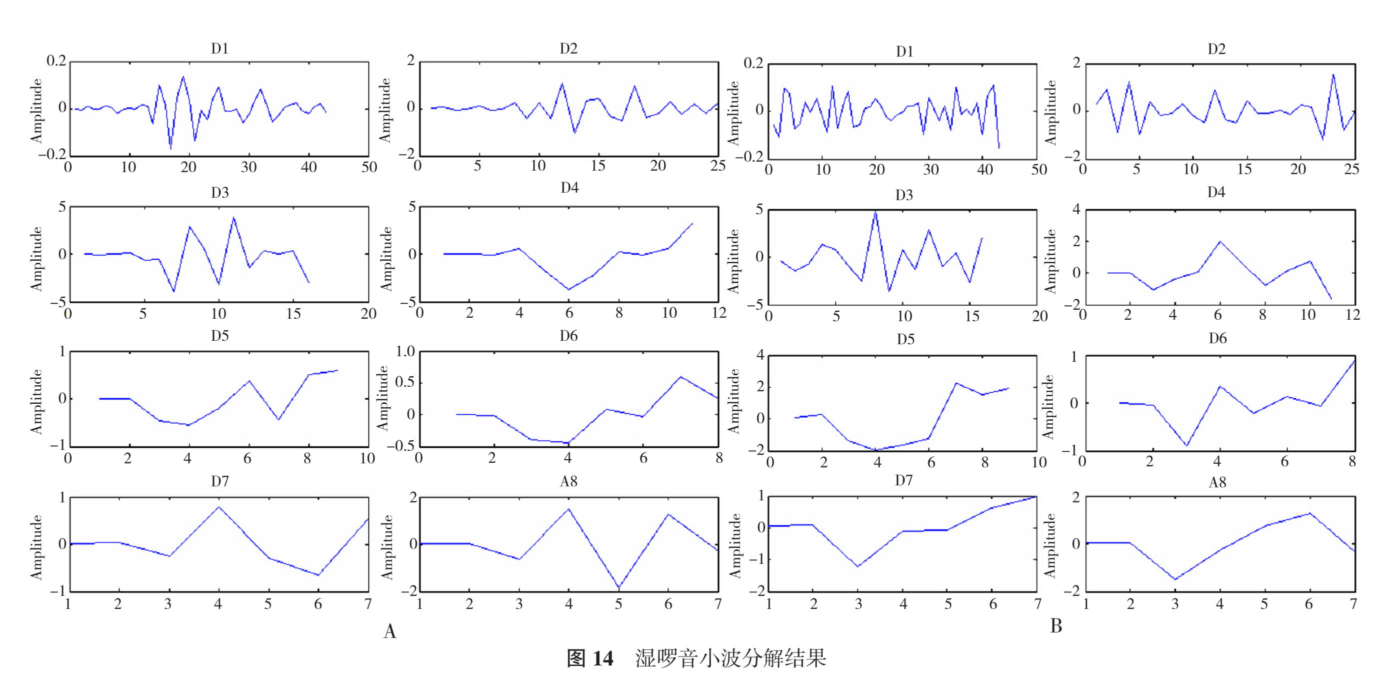


图 14 湿啰音小波分解结果

Figure 14 Results of the wavelet decomposition of moist rale

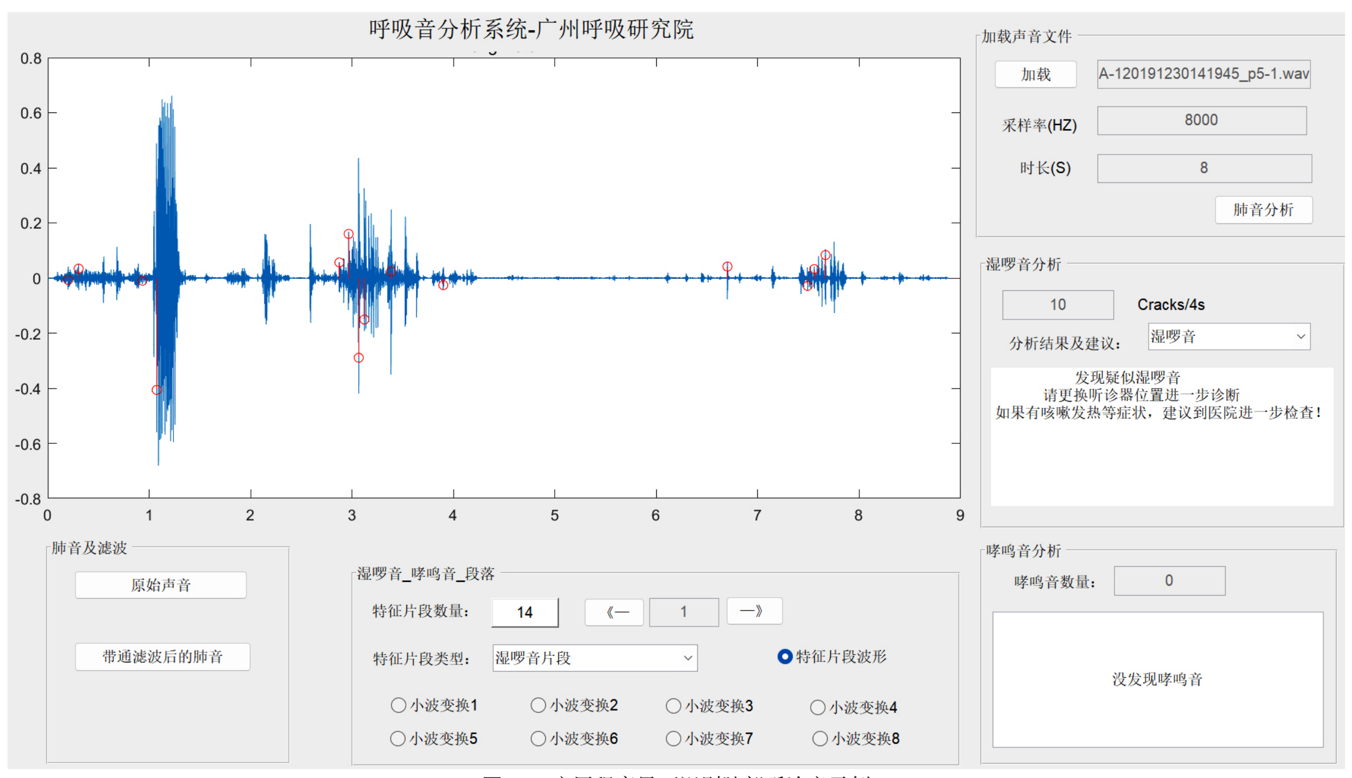


图 15 应用程序界面识别肺部听诊音示例

Figure 15 Example of an application interface to recognise lung auscultation sounds

的智能判读听诊器提供技术借鉴。

本研究局限性：目前这个数据库中的病理性肺音片段相对较少，这可能是由以下几个因素造成的，（1）本研究并没有涵盖那些病情特别严重的患者，这意味着那些可能会呈现更加明显肺部异常声音的患者群体并未被包括在内；（2）那些同时受多种呼吸系统疾病影响的患者也未能被纳入本研究，复杂的临床情况可能会造成多样化的肺音变化，而这部分信息对于肺音数据库而言是极其宝贵的；（3）医生在更新患者病历时可能会有所延迟，这可能导致在收集肺音的时候，患者的实际健康状况已经有所改善，比如，如果患者的病情在稳定期，那么其肺部的异常声音可能会相对较弱，甚至在检

测时不再明显。这种情况下，收集到的肺音可能就不能准确反映患者在更加活跃病程阶段的肺音特征。

致谢：上海拓萧智能科技有限公司协助算法开发。

作者贡献：张冬莹提出肺音研究思路，设计研究方案，研究命题的提出、设计，包括肺音听诊对象分组，伪湿罗音的智能识别和目标导向的研究思路（智能判读听诊器的推广应用导向），以及项目研究过程中专利申报的思路等负责论文起草，负责最终版本修订，对论文负责。张冬莹、李洽胜、梁振宇、简文华负责研究对象甄别入组，组织研究过程的实施，负责质量控制。张冬莹、叶培韬负责数据收集、采集、清洗和统计学分析、绘制图表等。郑劲平为研究项目顾问，指导研究方案及论文修改。

本文无利益冲突。

参考文献

[1] SRIVASTAVA A，JAIN S，MIRANDA R，et al. Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease [J]. PeerJ Comput Sci，2021，7：e369. DOI：10.7717/peerj-cs.369.

[2] BARDOU D，ZHANG K，AHMAD S M. Lung sounds classification using convolutional neural networks [J]. Artif Intell Med，2018，88：58-69. DOI：10.1016/j.artmed.2018.04.008.

[3] GOTTLIEB E R，ALIOTTA J M，TAMMARO D. Comparison of analogue and electronic stethoscopes for pulmonary auscultation by internal medicine residents [J]. Postgrad Med J，2018，94（1118）：700-703. DOI：10.1136/postgradmedj-2018-136052.

[4] KWON A M，KANG K. A temporal dependency feature in lower dimension for lung sound signal classification [J]. Sci Rep，2022，12：7889. DOI：10.1038/s41598-022-11726-3.

[5] CHAMBERLAIN D，KODGULE R，GANELIN D，et al. Application of semi-supervised deep learning to lung sound analysis [J]. Annu Int Conf IEEE Eng Med Biol Soc，2016，2016：804-807. DOI：10.1109/EMBC.2016.7590823.

[6] 叶培韬，张冬莹，陈文雅，等. 人工智能在呼吸音方面的研究进展 [J]. 国际呼吸杂志，2022，42（13）：961-966. DOI：10.3760/cma.j.cn131368-20220120-00045.

[7] AVILES-SOLIS J C，VANBELLE S，HALVORSEN P A，et al. International perception of lung sounds：a comparison of classification across some European borders [J]. BMJ Open Respir Res，2017，4（1）：e000250. DOI：10.1136/bmjresp-2017-000250.

[8] HOFFMANN C，FALZONE E，VERRET C，et al. Brief report：pulmonary auscultation in the operating room：a prospective randomized blinded trial comparing electronic and conventional stethoscopes [J]. Anesth Analg，2013，117（3）：646-648. DOI：10.1213/ANE.0b013e31829ec0a5.

[9] 上海市医学会儿科分会呼吸学组，福建省医学会儿科分会呼吸学组，江苏省医学会儿科分会呼吸学组，等. 高质量呼吸音数据库建设及管理规范专家共识 [J]. 中华实用儿科临床杂志，2023，38（7）：481-485. DOI：10.3760/cma.j.cn101070-20230120-00054.

[10] 姚小静，王洪，李燕，等. 肺音信号分析及其识别方法的研究进展 [J]. 重庆理工大学学报（自然科学），2013，27（12）：95-100.

[11] ZHANG K X，WANG X F，HAN F F，et al. The detection of crackles based on mathematical morphology in spectrogram analysis [J]. Technol Health Care，2015，23（Suppl 2）：S489-494. DOI：10.3233/THC-150986.

[12] REICHERT S，GASS R，BRANDT C，et al. Analysis of respiratory sounds：state of the art [J]. Clin Med Circ Respirat Pulm Med，2008，2：45-58. DOI：10.4137/ccrpm.s530.

[13] 宋倩，林和平，周琪栋. 支持向量机理论研究 [J]. 信息技术，2013，37（9）：152-154，159. DOI：10.3969/j.issn.1009-2552.2013.09.042.

[14] 刘国栋. 支持向量机在肺部病理音识别中的应用研究 [D]. 天津：南开大学，2015.

[15] YE P T，LI Q S，JIAN W H，et al. Regularity and mechanism of fake crackle noise in an electronic stethoscope [J]. Front Physiol，2022，13：1079468. DOI：10.3389/fphys.2022.1079468.

[16] 石陆魁，刘文浩，李站茹. 基于 LDA 和小波分解的肺音特征提取方法 [J]. 计算机工程与应用，2017，53（22）：116-120，149. DOI：10.3778/j.issn.1002-8331.1605-0330.

[17] 刘文青. 结合 EMD 和小波变换的肺音信号重构及呼吸周期分割研究 [D]. 武汉：江汉大学，2023.

（收稿日期：2024-01-22；修回日期：2024-04-10）

（本文编辑：毛亚敏）